

Supplementary Document of:

Diet-dependent acid load – the missing link between an animal protein-rich diet and non-alcoholic fatty liver disease?

Louise JM Alferink, Jessica C Kiefte-de Jong, Nicole S Erler, Robert J de Knegt, Ewout J Hoorn, M Arfan

Ikram, Harry LA Janssen, Herold J Metselaar, Oscar H Franco, Sarwa Darwish Murad

Supplementary Methods I: Detailed description of the study population

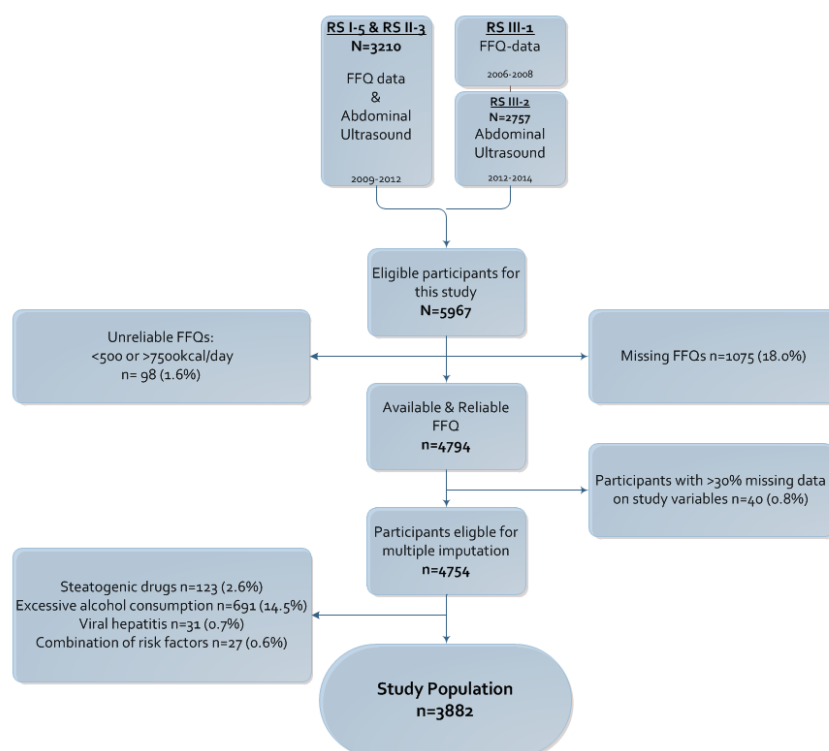


Figure is previously published in: Alferink LJM, Kieft-de Jong JC, Erler NS, et al. Association of dietary macronutrient composition and non-alcoholic fatty liver disease in an ageing population: the Rotterdam Study. Gut 2018. doi: 10.1136/gutjnl-2017-315940.

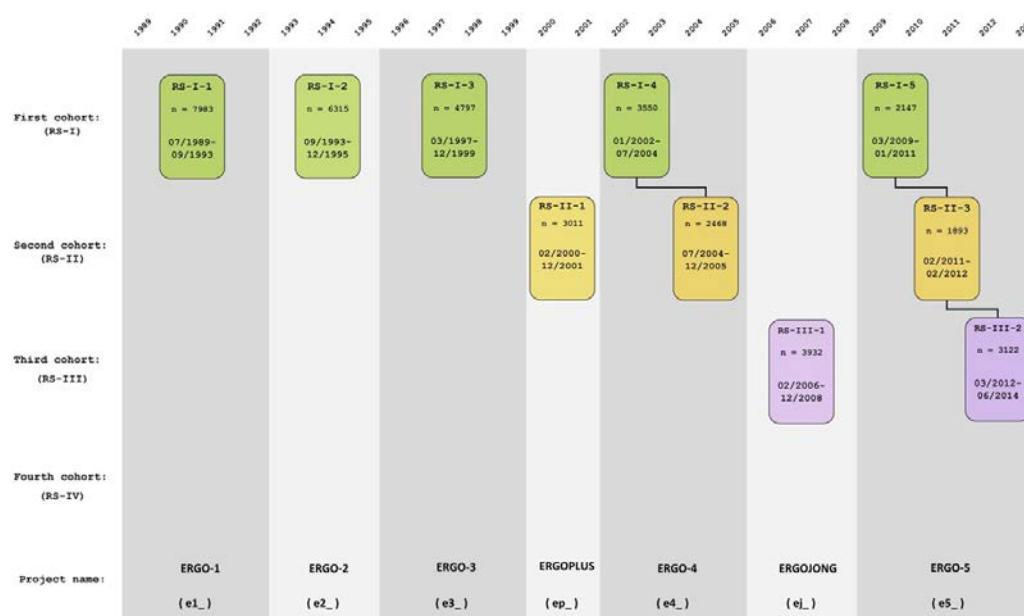


Diagram of examination cycles of the Rotterdam Study (RS). RS-I-1 refers to the baseline examination of the original cohort. RS-I-2, RS-I-3, RS-I-4, and RS-I-5 refer to re-examinations of the original cohort members. RS-II-1 refers to the extension of the cohort with persons from the study district that had become 55 years since the start of the study or those of 55 years or over that migrated into the study district. RS-II-2 and RS-II-3 refer to re-examinations of the extension cohort. RS-III-1 refers to the baseline examination of all persons aged 45 years and over living in the study district that had not been examined already (i.e., mainly comprising those aged 45–60 years). RS-III-2 refers to the first re-examination of this third cohort. Similarly, examinations RS-I-5, RS-II-3, and RS-III-2 share the same program items. Figure is previously published in: Ikram MA, Brusselle GGO, Murad SD, et al. The Rotterdam Study: 2018 update on objectives, design and main results. Eur J Epidemiol 2017;32(9):807-50.

Supplementary Methods II: details on the multiple imputation process

Variables (range missing values per variable: 0.02 to 10.79%) were imputed using multiple imputation under the fully conditioned specification to reduced bias due to missing data(1). Thirty imputed datasets were created using the R package mice(2), and analysed separately. Results from these analyses were pooled using Rubin's rules in order to take into account the added uncertainty due to the missing data(3).

	Multiple imputation
Software used	R version 3.5.0
Imputation method and key settings	Fully conditional specification (package mice version 3.1.0); maximum iterations: 20
No. of imputed data sets created	30
Analyses variables	total cholesterol; triglycerides; calcium intake; high density lipoprotein cholesterol; body mass index; glucose; weight; physical activity; smoking status; systolic blood pressure; diastolic blood pressure; ethnicity; education level; alcohol consumption; fat intake; waist circumference; kilocalorie intake; Dietary Quality; age; anti-diabetic drugs; lipid-lowering drugs; antihypertensive drugs; fibre intake; study cohort; gender; steatosis; protein intake; carbohydrate intake; vitamin E intake; potassium intake; magnesium intake; phosphorus intake; glomerular filtration rate
Auxiliary variables	aspartate transaminase; alanine transaminase; homeostasis model assessment of insulin resistance; hip circumference; heart rate; spleen size; creatinine; gamma-glutamyltransferase
Treatment of not normally distributed continuous variables	Predictive mean matching
Treatment of normally distributed variables	Linear regression
Treatment of binary/categorical variables	(Proportional odds) logistic regression
Population	For the imputation we used reliable and completed FFQs. In addition participants had to have less than 30% missing on study variables. Imputed population (n=4.754).

References:

1. Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *Bmj* 2009;338:b2393.
2. Buuren van S, Groothuis-Oudhoorn K. mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software, Articles* 2011;45:1-67.
3. Rubin DB. Multiple imputation for nonresponse in surveys. Wiley London 1987.

Supplementary Table 1: Imputation Characteristics

	Original Data n=3882	After imputation
Demographics		
Age (years)	69.7 (8.8)	no missing data
Female (%)	58.3	no missing data
Caucasian (%)	97.7	97.6
<u>Education Level (%)</u>		
Low	48.4	48.4
Intermediate	30.3	30.3
High	21.3	21.3
<u>Smoking status (%)</u>		
Never / Past or Current	36.7 / 63.3	36.1 / 63.9
Alcohol (units/d)	0.45 (0.05 – 1.19)	no missing data
Physical Activity (METh/wk)	41.3 (15.8 – 78.5)	40.5 (15.8 – 77.7)
Caloric Intake (kcal/day)	2031 (1621 – 2514)	no missing data
DQ	7 (6 – 8)	no missing data
PRAL	-4.7 (-15.4; 4.4)	no missing data
NEAP	35.7 (29.6 – 42.3)	no missing data
A:P	12.7 (10.2 – 15.4)	no missing data
Physical examination		
<u>BMI (kg/m²)</u>	26.9 (24.5 – 29.7)	26.9 (24.5 – 29.7)
Lean (%)	30.2	30.2
Overweight (%)	69.8	69.8
<u>WC (cm)</u>		
Men	98.2 (10.6)	98.2 (10.6)
Women	89.2 (12.2)	no missing data
Biochemistry		
AST (U/L)	24 (21 – 28)	24 (21 – 28)
ALT (U/L)	18 (15 – 24)	18 (15 – 24)
GGT (U/L)	23 (17 – 33)	23 (17 – 34)
Platelets (*10 ⁹ /L)	262 (223 – 305)	262 (223 – 305)
HOMA-IR	2.6 (1.7 – 4.1)	2.6 (1.7 – 4.1)
Total Cholesterol (mmol/L)	5.4 (1.1)	5.4 (1.1)
HDL-C (mmol/L)	1.5 (0.4)	1.5 (0.4)
Triglycerides (mmol/L)	1.3 (1.0 – 1.7)	1.3 (1.0 – 1.7)
GFR	76.1 (65.7 – 85.8)	76.1 (65.6 – 85.8)
Comorbidities		
<u>Metabolic Syndrome (%)</u>	53.2	53.2
- WC>88cm (♀) or >120cm(♂)	43.2	43.2
- Triglycerides >150mg/dL	46.2	46.0
- HDL-C <40mg/dL(♂) or 50mg/dL(♀)	44.8	44.7
- Blood pressure ≥130/85mmHg	84.3	84.3
- FG>100mg/dL	46.8	46.8
Diabetes Mellitus (%)	13.2	13.1
Hypertension (%)	74.0	74.0
NAFLD (%)	34.4	no missing data

Data is expressed as mean (SD), median (P25-P75) or percentage.

Abbreviations ALT: alanine aminotransferase; A:P: animal protein to potassium ratio; AST: aspartate aminotransferase; BMI: body mass index; DAL: dietary acid load; GFR: glomerular filtration rate; GGT: gamma-glutamyltransferase; HDL-C: high-density lipoprotein cholesterol; HOMA-IR: homeostasis model assessment of insulin resistance; NAFLD: non-alcoholic fatty liver disease; NEAP: net endogenous acid production; PRAL; potential renal acid load.

Supplementary Table 2: Characteristics of the study population

	No NAFLD n=2545 (65.6%)	NAFLD n=1337 (34.4%)	P- value*
Demographics			
Age (years)	69.6 (9.2)	70.0 (8.2)	0.180
Female (%)	59.1	56.8	0.195
Caucasian (%)	97.4	98.1	0.237
<u>Education Level (%)</u>			
Low	45.0	54.9	
Intermediate	31.4	28.2	<0.001
High	23.6	16.9	
<u>Smoking status (%)</u>			
Never	38.2	32.0	<0.001
Past or Current	61.8	68.0	
Alcohol (units/d)	0.45 (0.06 – 1.21)	0.43 (0.05 – 1.19)	0.422
Physical Activity (METh/wk)	43.7 (17.5 – 81.6)	34.6 (13.5 – 70.3)	<0.001
Energy intake	2052 (1642 – 2537)	1996 (1579 – 2456)	0.003
PRAL (mEq/d)	-5.5 (-16.1; 3.4)	-2.9 (-13.7; 6.1)	<0.001
NEAP (mEq/d)	35.1 (29.2 – 41.6)	37.0 (30.2 – 43.6)	<0.001
A:P (mEq/d)	12.4 (10.0 – 15.0)	13.3 (10.6 – 16.0)	<0.001
Physical examination			
<u>BMI (kg/m²)</u>	25.8 (23.7 – 28.1)	29.3 (27.0 – 32.3)	<0.001
lean (%)	40.9	9.9	<0.001
overweight (%)	59.1	90.1	<0.001
Waist Circumference (cm)	89.1 (11.1)	100.3 (11.4)	<0.001
Biochemistry			
AST (U/L)	24 (21 – 28)	25 (21 – 29)	<0.001
ALT (U/L)	17 (14 – 22)	21 (16 – 29)	<0.001
GGT (U/L)	21 (15 – 30)	28 (20 – 39)	<0.001
Platelets (*10 ⁹ /L)	260 (222 – 303)	266 (225 – 310)	0.053
HOMA-IR	2.1 (1.5 – 3.1)	4.1 (2.7 – 6.1)	<0.001
Total Cholesterol (mmol/L)	5.5 (1.1)	5.4 (1.1)	0.002
HDL-C (mmol/L)	1.5 (0.4)	1.3 (0.4)	<0.001
Triglycerides (mmol/L)	1.2 (0.9 – 1.5)	1.6 (1.2 – 2.1)	<0.001
<u>GFR</u>	76 (65 – 85)	76 (66 – 86)	0.272
<60/>60 (%)	15.7 / 84.3	14.9 / 85.1	0.544
Comorbidities			
<u>Metabolic Syndrome</u>	41.7	75.0	<0.001
- Waist Circumference	29.1	69.9	<0.001
- Triglycerides	38.9	59.4	<0.001
- HDL-Cholesterol	38.3	56.7	<0.001
- Blood pressure	80.5	91.4	<0.001
- Fasting Glucose	36.3	66.6	<0.001
Diabetes Mellitus (%)	7.5	23.7	<0.001
Hypertension (%)	68.9	83.7	<0.001

Pooled data based on 30 imputations represent % for categorical variables and for continuous variables mean (SD) or median (P25-P75). *P-value is based on ANOVA, Kruskal-Wallis test or Chi-square test and is the comparison between the no NAFLD and NAFLD columns.

Abbreviations ALT: alanine aminotransferase; A:P: animal protein to potassium ratio; AST: aspartate aminotransferase; BMI: body mass index; DAL: dietary acid load; GFR: glomerular filtration rate; GGT: gamma-glutamyltransferase; HDL-C: high-density lipoprotein cholesterol; HOMA-IR: homeostasis model assessment of insulin resistance; NAFLD: non-alcoholic fatty liver disease; NEAP: net endogenous acid production; PRAL; potential renal acid load

Supplementary Table 3A: Characteristics per quartile NEAP

NEAP					
	Quartile 1	Quartile 2	Quartile 3	Quartile 4	<i>P</i> -trend*
DAL					
PRAL	-24.1 (-33.6 ; -17.7)	-9.8 (-13.2 ; -6.8)	-0.33 (-2.5 ; 2.2)	10.9 (6.8 ; 17.6)	<0.001
NEAP	25.3 (21.9 – 27.8)	32.7 (31.1 – 34.1)	38.7 (37.2 – 40.4)	47.5 (44.6 – 53.0)	n/a
A:P	9.0 (7.5 – 10.4)	11.7 (10.5 – 13.1)	13.9 (12.5 – 15.3)	17.4 (15.3 – 20.0)	<0.001
Demographics					
Age	71.0 (8.7)	69.6 (8.5)	68.8 (8.8)	69.5 (9.2)	<0.001
Female (%)	68.5	60.8	54.0	49.9	<0.001
Caucasian (%)	97.8	98.2	97.9	96.7	0.160
<u>Education Level (%)</u>					
Low	49.9	49.1	45.2	49.6	0.022
Intermediate	28.2	27.8	33.7	31.5	
High	21.9	23.1	21.1	18.9	
<u>Smoking status (%)</u>					
Never	40.8	36.4	34.7	32.5	0.002
Current / Former	59.2	63.6	65.3	67.5	
Alcohol (units/d)	0.36 (0.03 – 1.08)	0.55 (0.09 – 1.25)	0.52 (0.09 – 1.25)	0.39 (0.03 – 1.16)	<0.001
Physical Activity (METH/wk)	43.2 (17.3 – 78.9)	47.1 (18.0 – 84.0)	38.1 (15.0 – 76.0)	33.3 (13.5 – 70.6)	<0.001
Energy intake (Kcal)	2019 (1613 – 2457)	2037 (1665 – 2491)	2066 (1626 – 2532)	2015 (1588 – 2551)	0.525
Physical examination					
BMI (kg/m ²)	26.6 (24.3 – 29.4)	26.6 (24.4 – 29.5)	27.0 (24.6 – 29.7)	27.3 (24.7 – 30.0)	0.012
Biochemistry					
AST (U/L)	24 (21 – 28)	24 (21 – 28)	24 (21 – 28)	25 (21 – 29)	0.357
ALT (U/L)	18 (14 – 23)	18 (14 – 24)	19 (15 – 24)	19 (15 – 25)	<0.001
GGT (U/L)	21 (16 – 30)	22 (17 – 32)	23 (17 – 34)	25 (18 – 37)	<0.001
Platelets (*10 ⁹ /L)	267 (232 – 312)	265 (220 – 303)	259 (220 – 300)	259 (220 – 306)	0.002
HOMA-IR	2.4 (1.7 – 3.6)	2.6 (1.7 – 4.0)	2.6 (1.8 – 4.1)	2.9 (1.9 – 4.7)	<0.001
Total Cholesterol (mmol/L)	5.6 (1.1)	5.5 (1.1)	5.4 (1.1)	5.3 (1.1)	<0.001
HDL-C (mmol/L)	1.5 (0.4)	1.5 (0.4)	1.5 (0.4)	1.4 (0.4)	<0.001
Triglycerides (mmol/L)	1.3 (1.0 – 1.7)	1.2 (1.0 – 1.7)	1.3 (1.0 – 1.7)	1.3 (1.0 – 1.8)	0.708
GFR	76 (65 – 85)	77 (66 – 86)	76 (66 – 86)	76 (65 – 86)	0.898
Comorbidities					
<u>Metabolic Syndrome</u>	49.9	51.1	54.9	56.6	0.009
- Waist Circumference	40.6	42.6	43.8	45.8	0.131
- Triglycerides	43.1	44.5	48.3	48.1	0.045
- HDL-Cholesterol	41.9	43.2	46.0	47.5	0.050
- Blood pressure	86.2	83.8	82.5	84.6	0.150
- Fasting Glucose	40.4	46.6	46.6	53.5	<0.001
Diabetes Mellitus (%)	11.2	12.5	11.7	17.1	<0.001
Hypertension (%)	75.9	72.9	71.8	75.3	0.131
NAFLD	31.0	31.5	35.8	39.4	<0.001

Data is expressed as mean (SD), median (P25-P75) or percentage. **P*-value is based on ANOVA, Kruskal-Wallis test or Chi-square test.

Abbreviations ALT: alanine aminotransferase; A:P: animal protein to potassium ratio; AST: aspartate aminotransferase; BMI: body mass index; DAL: dietary acid load; GFR: glomerular filtration rate; GGT: gamma-glutamyltransferase; HDL-C: high-density lipoprotein cholesterol; HOMA-IR: homeostasis model assessment of insulin resistance; NAFLD: non-alcoholic fatty liver disease; NEAP: net endogenous acid production; PRAL: potential renal acid load.

Supplementary Table 3B: Characteristics per quartile AP-ratio

A:P					
	Quartile 1	Quartile 2	Quartile 3	Quartile 4	P-trend*
DAL					
PRAL	-21.4 (-32.4 ; -13.0)	-10.4 (-16.4 ; -3.9)	-1.7 (-6.8 ; 3.1)	9.2 (3.5 ; 16.4)	<0.001
NEAP	26.1 (21.9 – 30.2)	32.4 (29.5 – 36.0)	37.7 (34.6 – 41.1)	46.2 (42.2 – 52.1)	<0.001
A:P	8.5 (7.2 – 9.5)	11.4 (10.8 – 12.0)	13.9 (13.3 – 14.6)	17.8 (16.4 – 20.1)	n/a
Demographics					
Age	69.9 (8.9)	69.2 (8.6)	69.4 (9.0)	70.4 (8.9)	0.009
Female (%)	63.5	57.3	57.1	55.3	0.001
Caucasian (%)	96.6	98.3	97.6	97.9	0.116
<u>Education Level (%)</u>					
Low	49.0	45.2	45.6	53.9	<0.001
Intermediate	28.1	31.7	31.2	30.2	
High	22.9	23.1	23.2	15.9	
<u>Smoking status (%)</u>					
Never	40.1	38.2	35.7	30.4	<0.001
Current / Former	59.9	61.8	64.3	69.6	
Alcohol (units/d)	0.38 (0.03 – 1.11)	0.52 (0.09 – 1.22)	0.45 (0.07 – 1.24)	0.45 (0.04 – 1.18)	0.002
Physical Activity (METh/wk)	45.0 (18.0 – 82.8)	44.4 (18.0 – 82.2)	36.8 (14.5 – 74.5)	32.2 (13.5 – 73.2)	<0.001
Energy intake (kcal)	2099 (1702 – 2598)	2063 (1686 – 2505)	2022 (1611 – 2526)	1940 (1504 – 2405)	<0.001
Physical Examination					
BMI (kg/m ²)	26.0 (23.9 – 28.9)	26.9 (24.5 – 29.4)	27.0 (24.8 – 29.8)	27.5 (24.9 – 30.5)	<0.001
Biochemistry					
AST (U/L)	24 (21 – 28)	24 (21 – 28)	24 (21 – 29)	24 (21 – 28)	0.605
ALT (U/L)	18 (14 – 23)	18 (14 – 23)	19 (15 – 24)	19 (15 – 25)	0.004
GGT (U/L)	21 (16 – 31)	22 (16 – 31)	24 (17 – 35)	25 (18 – 36)	<0.001
Platelets (*10 ⁹ /L)	268 (228 – 308)	261 (225 – 304)	259 (219 – 304)	261 (221 – 305)	0.054
HOMA-IR	2.3 (1.6 – 3.5)	2.5 (1.8 – 3.8)	2.7 (1.8 – 4.2)	2.9 (1.8 – 4.7)	<0.001
Total Cholesterol (mmol/L)	5.5 (1.1)	5.5 (1.1)	5.4 (1.1)	5.3 (1.1)	<0.001
HDL-C (mmol/L)	1.5 (0.4)	1.5 (0.4)	1.5 (0.4)	1.4 (0.4)	0.012
Triglycerides (mmol/L)	1.2 (1.0 – 1.7)	1.3 (1.0 – 1.7)	1.3 (1.0 – 1.8)	1.3 (1.0 – 1.8)	0.053
GFR	77 (66 – 86)	77 (66 – 85)	75 (66 – 86)	75 (64 – 85)	0.241
Comorbidities					
<u>Metabolic Syndrome</u>	45.3	52.0	55.6	59.7	<0.001
- Waist Circumference	35.0	41.4	45.7	50.5	<0.001
- Triglycerides	40.9	45.3	48.1	49.7	<0.001
- HDL-Cholesterol	39.5	44.7	46.5	48.0	0.001
- Blood pressure	83.4	83.0	84.9	85.9	0.276
- Fasting Glucose	37.7	46.4	49.7	53.3	<0.001
Diabetes Mellitus (%)	10.1	12.1	13.6	16.6	<0.001
Hypertension (%)	71.4	71.5	76.0	77.0	0.005
NAFLD	28.9	31.8	35.7	41.3	<0.001

Data is expressed as mean (SD), median (P25-P75) or percentage. *P-value is based on ANOVA, Kruskal-Wallis test or Chi-square test.

Abbreviations ALT: alanine aminotransferase; A:P: animal protein to potassium ratio; AST: aspartate aminotransferase; BMI: body mass index; DAL: dietary acid load; GFR: glomerular filtration rate; GGT: gamma-glutamyltransferase; HDL-C: high-

density lipoprotein cholesterol; HOMA-IR: homeostasis model assessment of insulin resistance; NAFLD: non-alcoholic fatty liver disease; NEAP: net endogenous acid production; PRAL; potential renal acid load.

Supplementary Table 4: Logistic regression analyses of DAL-residuals with outcome variable NAFLD

Total population (n=3882)			
	Q2 (n=970)	Q3 (n=971)	Q4 (n=970)
PRAL n per quartile	(n=970)	(n=971)	(n=970)
NEAP n per quartile	(n=970)	(n=971)	(n=970)
A:P n per quartile	(n=970)	(n=971)	(n=970)
Model 1 (sociodemographic)			
PRAL	0.98 (0.81 – 1.20)	1.16 (0.96 – 1.41)	1.37 (1.13 – 1.66)†
NEAP	1.03 (0.85 – 1.25)	1.25 (1.04 – 1.52)†	1.40 (1.16 – 1.69)†
A:P	1.16 (0.96 – 1.41)	1.37 (1.13 – 1.66)†	1.61 (1.33 – 1.95)†
Model 2 (lifestyle)			
PRAL	0.98 (0.81 – 1.19)	1.14 (0.94 – 1.39)	1.33 (1.10 – 1.62)†
NEAP	1.03 (0.85 – 1.25)	1.24 (1.02 – 1.50)	1.36 (1.12 – 1.65)†
A:P	1.16 (0.95 – 1.41)	1.33 (1.10 – 1.62)†	1.56 (1.29 – 1.89)†
Model 3 (metabolic)			
PRAL	0.97 (0.78 – 1.22)	1.10 (0.88 – 1.38)	1.23 (0.98 – 1.54)
NEAP	0.98 (0.78 – 1.22)	1.19 (0.96 – 1.49)	1.25 (1.00 – 1.56)
A:P	1.01 (0.81 – 1.27)	1.13 (0.90 – 1.41)	1.23 (0.98 – 1.53)
Model 4 (metabolic + DQ)			
PRAL	0.97 (0.77 – 1.22)	1.10 (0.88 – 1.38)	1.23 (0.97 – 1.55)
NEAP	0.98 (0.78 – 1.23)	1.20 (0.96 – 1.50)	1.25 (1.00 – 1.58)
A:P	1.01 (0.81 – 1.27)	1.13 (0.90 – 1.41)	1.23 (0.97 – 1.55)

Values are odds ratios with 95% confidence intervals taking quartile 1 as reference. **Bold** values indicate $P < 0.05$. † Indicates significant values using $P < 0.028$ as determined by Sidák.

Model 1 (socio-demographic) is adjusted for age, gender, education level, energy intake and study cohort **Model 2** (lifestyle) is in addition previous model adjusted for past or current smoking, units of alcohol, and physical activity **Model 3** (metabolic) is in addition to the previous model adjusted for HDL-cholesterol, triglycerides, metabolic syndrome, GFR, diabetes mellitus and log-transformed BMI **Model 4** (metabolic + DQ) is in addition to the previous model adjusted for DQ.

Abbreviations A:P: animal protein to potassium ratio; BMI: body mass index; DAL: dietary acid load; DQ: dietary quality; GFR: glomerular filtration rate; NAFLD: non-alcoholic fatty liver disease; NEAP: net endogenous acid production; PRAL; potential renal acid load.

Supplementary Table 5: Logistic regression analyses of DAL associations with outcome variable NAFLD stratified by sex

Men (n=1619)			
	Q2 (n=364)	Q3 (n=438)	Q4 (n=502)
PRAL n per quartile			
NEAP n per quartile	(n=380)	(n=447)	(n=486)
A:P n per quartile	(n=414)	(n=417)	(n=434)
Model 2 (lifestyle)			
PRAL	0.95 (0.68 – 1.33)	1.20 (0.87 – 1.64)	1.42 (1.05 – 1.93)†
NEAP	1.12 (0.80 – 1.55)	1.34 (0.98 – 1.85)	1.49 (1.09 – 2.04)†
A:P	1.35 (0.98 – 1.85)	1.59 (1.16 – 2.18)†	1.87 (1.37 – 2.55)†
Model 3 (metabolic)			
PRAL	0.90 (0.61 – 1.33)	1.14 (0.79 – 1.63)	1.30 (0.92 – 1.85)
NEAP	0.97 (0.66 – 1.42)	1.35 (0.94 – 1.94)	1.34 (0.93 – 1.91)
A:P	1.06 (0.74 – 1.52)	1.31 (0.92 – 1.87)	1.45 (1.02 – 2.07)
Women (n=2263)			
	Q2 (n=606)	Q3 (n=533)	Q4 (n=468)
PRAL n per quartile			
NEAP n per quartile	(n=590)	(n=524)	(n=484)
A:P n per quartile	(n=556)	(n=554)	(n=536)
Model 2 (lifestyle)			
PRAL	1.02 (0.80 – 1.30)	1.05 (0.81 – 1.34)	1.34 (1.04 – 1.72)†
NEAP	1.00 (0.79 – 1.28)	1.15 (0.90 – 1.47)	1.28 (1.00 – 1.65)
A:P	1.05 (0.82 – 1.36)	1.20 (0.93 – 1.54)	1.42 (1.11 – 1.82)†
Model 3 (metabolic)			
PRAL	0.99 (0.75 – 1.32)	1.04 (0.78 – 1.40)	1.22 (0.90 – 1.65)
NEAP	0.98 (0.74 – 1.30)	1.08 (0.81 – 1.45)	1.17 (0.87 – 1.58)
A:P	0.91 (0.68 – 1.22)	0.94 (0.70 – 1.26)	1.08 (0.81 – 1.46)

Values are odds ratios with 95% confidence intervals taking quartile 1 as reference. *P*-trend is calculated across the quartiles. **Bold** values indicate $P < 0.05$. † Indicates significant values using $P < 0.028$ as determined by Sidák.

Model 1 (socio-demographic) is adjusted for age, education level, energy intake and study cohort **Model 2** (lifestyle) is in addition previous model adjusted for past or current smoking, units of alcohol, and physical activity **Model 3** (metabolic) is in addition to the previous model adjusted for HDL-cholesterol, triglycerides, metabolic syndrome, GFR, diabetes mellitus and BMI **Model 4** (metabolic + DQ) is in addition to the previous model adjust for DQ.

Abbreviations A:P: animal protein to potassium ratio; BMI: body mass index; DAL: dietary acid load; DQ: dietary quality; GFR: glomerular filtration rate; NAFLD: non-alcoholic fatty liver disease; NEAP: net endogenous acid production; PRAL: potential renal acid load.

Supplementary Table 6: Logistic regression analyses of DAL associations with outcome variable NAFLD stratified by study cohort

RS-cohort I (n=1069)			
	Q2	Q3	Q4
PRAL n per quartile	(n=259)	(n=259)	(n=263)
NEAP n per quartile	(n=259)	(n=234)	(n=263)
A:P n per quartile	(n=239)	(n=262)	(n=283)
Model 2 (lifestyle)			
PRAL	0.93 (0.64 – 1.34)	0.92 (0.63 – 1.34)	1.07 (0.74 – 1.54)
NEAP	1.10 (0.77 – 1.57)	1.08 (0.74 – 1.55)	1.10 (0.77 – 1.57)
A:P	1.21 (0.83 – 1.76)	1.11 (0.77 – 1.60)	1.45 (1.02 – 2.07)
Model 3 (metabolic)			
PRAL	1.03 (0.67 – 1.58)	1.00 (0.64 – 1.55)	1.19 (0.77 – 1.83)
NEAP	1.13 (0.74 – 1.72)	1.14 (0.73 – 1.76)	1.22 (0.80 – 1.86)
A:P	1.22 (0.78 – 1.90)	1.05 (0.68 – 1.62)	1.40 (0.92 – 2.13)
RS-cohort II (n=1183)			
	Q2	Q3	Q4
PRAL n per quartile	(n=303)	(n=315)	(n=270)
NEAP n per quartile	(n=284)	(n=295)	(n=289)
A:P n per quartile	(n=297)	(n=281)	(n=316)
Model 2 (lifestyle)			
PRAL	1.00 (0.70 – 1.42)	1.32 (0.93 – 1.87)	1.73 (1.21 – 2.46)†
NEAP	0.99 (0.69 – 1.40)	1.31 (0.93 – 1.85)	1.65 (1.17 – 2.32)†
A:P	1.10 (0.77 – 1.57)	1.61 (1.13 – 2.29)†	1.70 (1.20 – 2.40)†
Model 3 (metabolic)			
PRAL	0.75 (0.49 – 1.15)	1.20 (0.80 – 1.82)	1.43 (0.94 – 2.19)
NEAP	0.76 (0.50 – 1.16)	1.14 (0.76 – 1.70)	1.43 (0.95 – 2.16)
A:P	0.77 (0.51 – 1.18)	1.22 (0.80 – 1.86)	1.11 (0.73 – 1.68)
RS-cohort III (n=1630)			
	Q2	Q3	Q4
PRAL n per quartile	(n=408)	(n=397)	(n=437)
NEAP n per quartile	(n=427)	(n=442)	(n=418)
A:P n per quartile	(n=434)	(n=428)	(n=371)
Model 2 (lifestyle)			
PRAL	1.05 (0.77 – 1.44)	1.10 (0.80 – 1.51)	1.44 (1.07 – 1.95)†
NEAP	1.00 (0.73 – 1.38)	1.24 (0.90 – 1.69)	1.40 (1.02 – 1.92)
A:P	1.12 (0.83 – 1.52)	1.30 (0.96 – 1.77)	1.61 (1.18 – 2.19)†
Model 3 (metabolic)			
PRAL	1.12 (0.79 – 1.60)	1.03 (0.73 – 1.47)	1.24 (0.88 – 1.76)
NEAP	1.01 (0.71 – 1.45)	1.19 (0.83 – 1.69)	1.16 (0.81 – 1.67)
A:P	0.94 (0.66 – 1.33)	1.03 (0.73 – 1.45)	1.18 (0.83 – 1.69)

Values are odds ratios with 95% confidence intervals taking quartile 1 as reference. **Bold** values indicate $P < 0.05$. † Indicates significant values using $P < 0.028$ as determined by Sidák.

Model 1 (socio-demographic) is adjusted for age, gender, education level and energy intake **Model 2** (lifestyle) is in addition previous model adjusted for past or current smoking, units of alcohol, and physical activity **Model 3** (metabolic) is in addition to the previous model adjusted for HDL-cholesterol, triglycerides, metabolic syndrome, GFR, diabetes mellitus and BMI **Model 4** (metabolic + DQ) is in addition to the previous model adjust for DQ.

Abbreviations A:P: animal protein to potassium ratio; BMI: body mass index; DAL: dietary acid load; DQ: dietary quality; GFR: glomerular filtration rate; NAFLD: non-alcoholic fatty liver disease; NEAP: net endogenous acid production; PRAL: potential renal acid load.

Supplementary Table 7: Logistic regression analyses of DAL associations with outcome variable NAFLD stratified by GFR (cut-off 60)

GFR < 60 (n=597)			
	Q2 (n=147)	Q3 (n=143)	Q4 (n=162)
PRAL n per quartile			
NEAP n per quartile	(n=142)	(n=149)	(n=156)
A:P n per quartile	(n=144)	(n=144)	(n=174)
Model 2 (lifestyle)			
PRAL	1.00 (0.59 – 1.69)	0.98 (0.58 – 1.68)	1.49 (0.90 – 2.47)
NEAP	0.89 (0.53 – 1.51)	1.09 (0.66 – 1.82)	1.23 (0.75 – 2.02)
A:P	1.08 (0.63 – 1.86)	1.33 (0.78 – 2.26)	1.66 (1.00 – 2.74)
Model 3 (metabolic)			
PRAL	1.05 (0.57 – 1.91)	1.01 (0.55 – 1.87)	1.81 (1.01 – 3.24)
NEAP	0.86 (0.47 – 1.56)	1.25 (0.70 – 2.23)	1.40 (0.79 – 2.49)
A:P	1.11 (0.60 – 2.05)	1.39 (0.75 – 2.55)	1.68 (0.94 – 2.99)
GFR > 60 (n=3285)			
	Q2 (n=823)	Q3 (n=828)	Q4 (n=808)
PRAL n per quartile			
NEAP n per quartile	(n=828)	(n=822)	(n=814)
A:P n per quartile	(n=826)	(n=827)	(n=796)
Model 2 (lifestyle)			
PRAL	0.99 (0.80 – 1.22)	1.14 (0.92 – 1.41)	1.37 (1.11 – 1.69)†
NEAP	1.05 (0.85 – 1.30)	1.25 (1.02 – 1.55)	1.38 (1.12 – 1.70)†
A:P	1.16 (0.94 – 1.43)	1.34 (1.08 – 1.65)†	1.56 (1.27 – 1.93)†
Model 3 (metabolic)			
PRAL	0.96 (0.75 – 1.23)	1.09 (0.86 – 1.39)	1.18 (0.92 – 1.50)
NEAP	0.99 (0.78 – 1.27)	1.16 (0.91 – 1.48)	1.20 (0.94 – 1.53)
A:P	0.93 (0.73 – 1.18)	1.04 (0.82 – 1.33)	1.13 (0.88 – 1.44)

Values are odds ratios with 95% confidence intervals taking quartile 1 as reference. *P*-trend is calculated across the quartiles. **Bold** values indicate $P < 0.05$. † Indicates significant values using $P < 0.028$ as determined by Sidák.

Model 1 (socio-demographic) is adjusted for age, gender, education level, energy intake and study cohort **Model 2** (lifestyle) is in addition previous model adjusted for past or current smoking, units of alcohol, and physical activity **Model 3** (metabolic) is in addition to the previous model adjusted for HDL-cholesterol, triglycerides, metabolic syndrome, diabetes mellitus and BMI **Model 4** (metabolic + DQ) is in addition to the previous model adjust for DQ.

Abbreviations A:P: animal protein to potassium ratio; BMI: body mass index; DAL: dietary acid load; DQ: dietary quality; GFR: glomerular filtration rate; NAFLD: non-alcoholic fatty liver disease; NEAP: net endogenous acid production; PRAL: potential renal acid load.

Supplementary Table 8: Logistic regression analyses of DAL associations with outcome variable NAFLD stratified by age (cut-off 65 years)

Age <65 years old (n=1160)			
	Q2 (n=281)	Q3 (n=276)	Q4 (n=333)
PRAL n per quartile	(n=281)	(n=276)	(n=333)
NEAP n per quartile	(n=289)	(n=318)	(n=319)
A:P n per quartile	(n=301)	(n=311)	(n=257)
Model 2 (lifestyle)			
PRAL	1.14 (0.78 – 1.68)	1.32 (0.90 – 1.94)	1.58 (1.10 – 2.28)†
NEAP	1.06 (0.71 – 1.57)	1.33 (0.91 – 1.96)	1.47 (1.00 – 2.16)
A:P	1.14 (0.78 – 1.64)	1.33 (0.93 – 1.92)	1.63 (1.13 – 2.37)†
Model 3 (metabolic)			
PRAL	1.28 (0.83 – 1.98)	1.29 (0.83 – 1.98)	1.39 (0.92 – 2.12)
NEAP	1.12 (0.71 – 1.76)	1.36 (0.88 – 2.09)	1.29 (0.83 – 2.00)
A:P	0.99 (0.65 – 1.50)	1.100 (0.73 – 1.66)	1.23 (0.80 – 1.88)
Age ≥ 65 years old (n=2722)			
	Q2 (n=689)	Q3 (n=695)	Q4 (n=637)
PRAL n per quartile	(n=689)	(n=695)	(n=637)
NEAP n per quartile	(n=681)	(n=653)	(n=651)
A:P n per quartile	(n=669)	(n=660)	(n=713)
Model 2 (lifestyle)			
PRAL	0.94 (0.75 – 1.18)	1.05 (0.83 – 1.32)	1.34 (1.07 – 1.69)†
NEAP	1.02 (0.81 – 1.28)	1.19 (0.95 – 1.49)	1.34 (1.07 – 1.68)†
A:P	1.14 (0.90 – 1.44)	1.33 (1.06 – 1.68)†	1.56 (1.24 – 1.95)†
Model 3 (metabolic)			
PRAL	0.87 (0.66 – 1.13)	1.01 (0.77 – 1.32)	1.26 (0.96 – 1.65)
NEAP	0.92 (0.71 – 1.20)	1.10 (0.84 – 1.43)	1.27 (0.97 – 1.65)
A:P	0.94(0.72 – 1.23)	1.10 (0.84 – 1.44)	1.21 (0.93 – 1.58)

Values are odds ratios with 95% confidence intervals taking quartile 1 as reference. *P*-trend is calculated across the quartiles. **Bold** values indicate $P < 0.05$. † Indicates significant values using $P < 0.028$ as determined by Sidák.

Model 1 (socio-demographic) is adjusted for gender, education level, energy intake and study cohort **Model 2** (lifestyle) is in addition previous model adjusted for past or current smoking, units of alcohol, and physical activity **Model 3** (metabolic) is in addition to the previous model adjusted for HDL-cholesterol, triglycerides, metabolic syndrome, GFR, diabetes mellitus and BMI **Model 4** (metabolic + DQ) is in addition to the previous model adjust for DQ.

Abbreviations A:P: animal protein to potassium ratio; BMI: body mass index; DAL: dietary acid load; DQ: dietary quality; GFR: glomerular filtration rate; NAFLD: non-alcoholic fatty liver disease; NEAP: net endogenous acid production; PRAL; potential renal acid load.

Supplementary Table 9: Logistic regression analyses of DAL associations with outcome variable NAFLD stratified by BMI (cut-off 25)

BMI < 25 kg/m² (n=1174)			
	Q2	Q3	Q4
PRAL n per quartile	(n=296)	(n=285)	(n=267)
NEAP n per quartile	(n=308)	(n=273)	(n=277)
A:P n per quartile	(n=289)	(n=259)	(n=251)
Model 2 (lifestyle)			
PRAL	1.20 (0.71 – 2.05)	1.11 (0.64 – 1.92)	1.53 (0.91 – 2.58)
NEAP	1.41 (0.82 – 2.41)	1.30 (0.74 – 2.29)	1.70 (0.99 – 2.90)
A:P	1.32 (0.79 – 2.19)	1.18 (0.69 – 2.02)	1.63 (0.98 – 2.70)
Model 3 (metabolic)			
PRAL	1.15 (0.67 – 1.99)	1.00 (0.57 – 1.78)	1.62 (0.95 – 2.78)
NEAP	1.35 (0.78 – 2.36)	1.18 (0.66 – 2.12)	1.82 (1.05 – 3.18)
A:P	1.19 (0.70 – 2.03)	1.08 (0.62 – 1.89)	1.60 (0.95 – 2.70)
BMI ≥ 25 kg/m² (n=2708)			
	Q2	Q3	Q4
PRAL n per quartile	(n=674)	(n=686)	(n=703)
NEAP n per quartile	(n=662)	(n=698)	(n=693)
A:P n per quartile	(n=681)	(n=712)	(n=719)
Model 2 (lifestyle)			
PRAL	0.91 (0.73 – 1.14)	1.10 (0.88 – 1.38)	1.32 (1.06 – 1.65)†
NEAP	0.95 (0.76 – 1.19)	1.16 (0.93 – 1.45)	1.31 (1.05 – 1.64)†
A:P	0.94 (0.75 – 1.18)	1.13 (0.90 – 1.41)	1.34 (1.07 – 1.68)†
Model 3 (metabolic)			
PRAL	0.88 (0.69 – 1.13)	1.11 (0.87 – 1.41)	1.22 (0.96 – 1.55)
NEAP	0.92 (0.72 – 1.17)	1.21 (0.96 – 1.53)	1.20 (0.96 – 1.53)
A:P	0.92 (0.72 – 1.17)	1.10 (0.86 – 1.39)	1.25 (0.98 – 1.59)

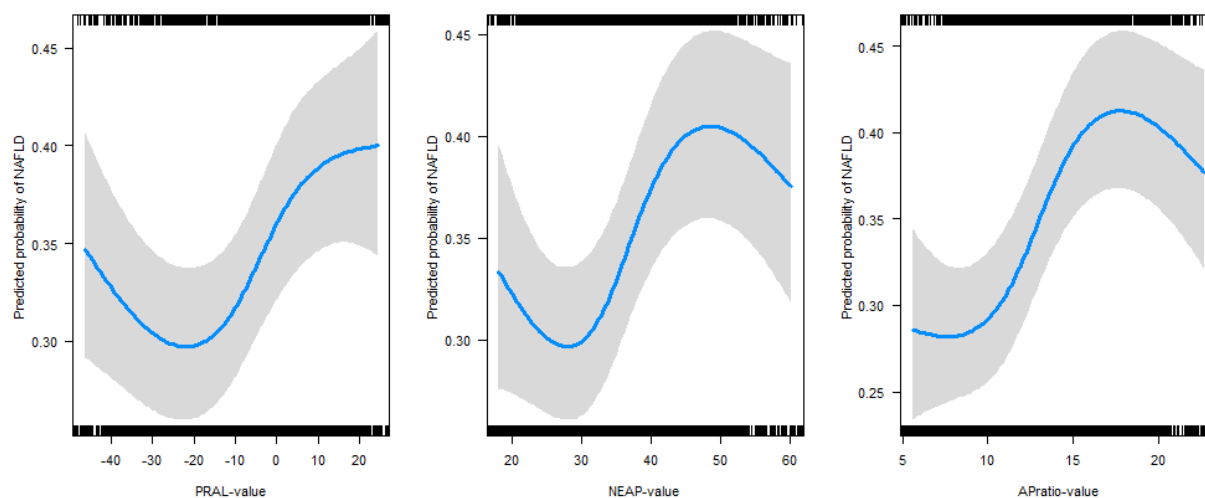
Values are odds ratios with 95% confidence intervals taking quartile 1 as reference. *P*-trend is calculated across the quartiles.

Model 1 (socio-demographic) is adjusted for age, gender, education level, energy intake and study cohort **Model 2** (lifestyle) is in addition previous model adjusted for past or current smoking, units of alcohol, and physical activity **Model 3** (metabolic) is in addition to the previous model adjusted for HDL-cholesterol, triglycerides metabolic syndrome, GFR and diabetes mellitus **Model 4** (metabolic + DQ) is in addition to the previous model adjust for DQ.

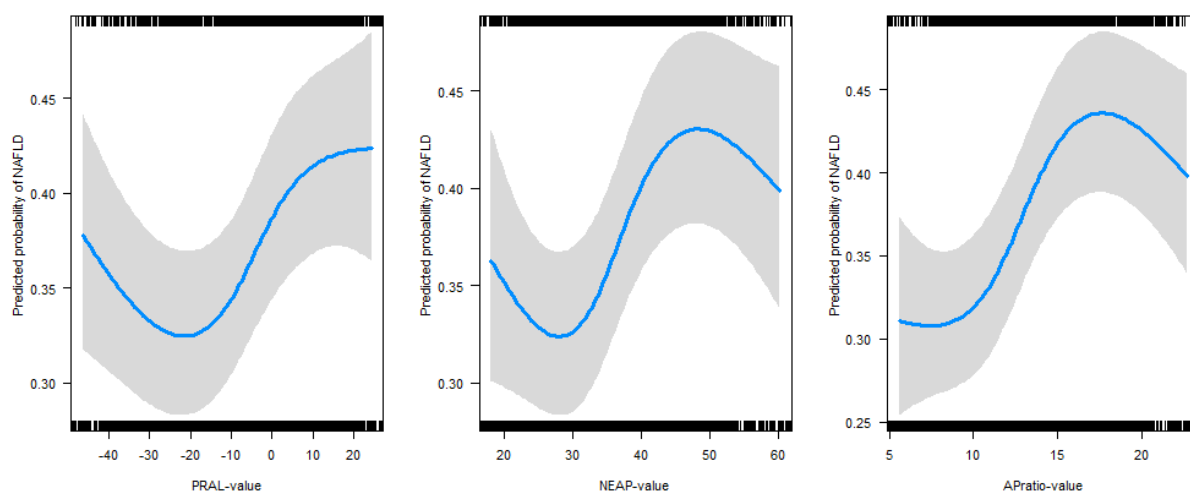
Abbreviations A:P: animal protein to potassium ratio; BMI: body mass index; DAL: dietary acid load; DQ: dietary quality; GFR: glomerular filtration rate; NAFLD: non-alcoholic fatty liver disease; NEAP: net endogenous acid production; PRAL; potential renal acid load.

Supplementary Figure 1: Natural cubic splines for the association of DAL-proxies with predicted probability of NAFLD in model 1, 2, and 4.

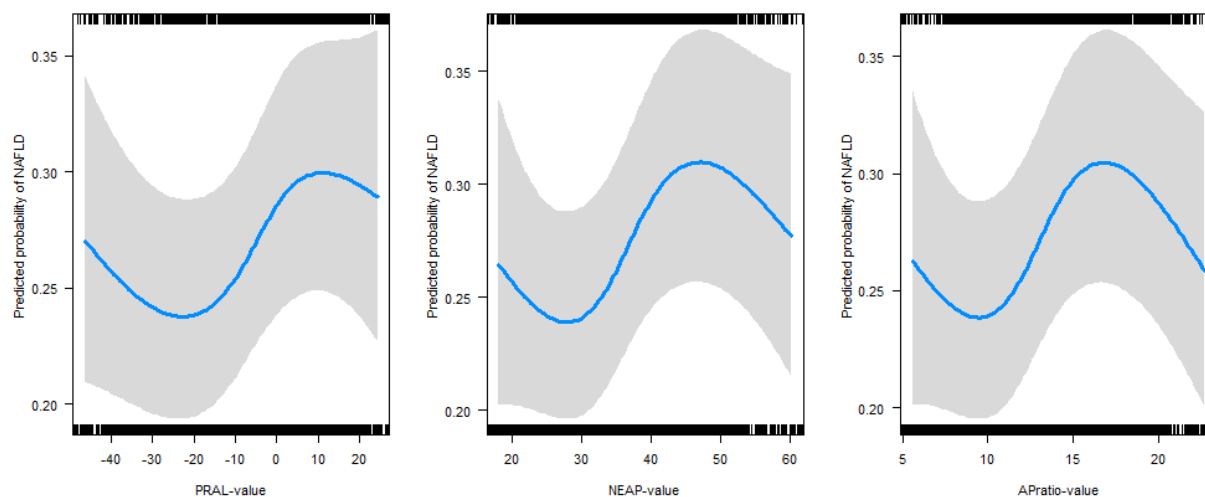
A. (Sociodemographic) model 1



B. (Lifestyle) model 2



C. (Metabolic + DQ) model 4



Model 1 (socio-demographic) is adjusted for age, gender, education level, energy intake and study cohort

Model 2 (lifestyle) is in addition previous model adjusted for past or current smoking, units of alcohol, and physical activity

Model 4 (metabolic + DQ) is in addition to model 2 adjusted for BMI, HDL-cholesterol, triglycerides, metabolic syndrome, GFR, diabetes mellitus, and DQ.

Abbreviations A:P: animal protein to potassium ratio; BMI: body mass index; DAL: dietary acid load; DQ: dietary quality; GFR: glomerular filtration rate; NAFLD: non-alcoholic fatty liver disease; NEAP: net endogenous acid production; PRAL: potential renal acid load.

Comparison model with DAL-proxy vs model without DAL-proxy

A. (Sociodemographic) model 1 (PRAL: $P=7.4e^{-5}$, NEAP: $P=1.2e^{-5}$, and A:P: $P=1.9e^{-8}$).

B. (Lifestyle) model 2 (PRAL: $P=2.3e^{-4}$, NEAP: $P=4.0e^{-5}$, and A:P: $P=1.3e^{-7}$)

C. (Metabolic + DQ) model 4 (PRAL: $P=0.043$, NEAP: $P=0.017$, and A:P: $P=0.014$)